encyclopedia.com

William Sealy Gosset | Encyclopedia.com

Complete Dictionary of Scientific Biography
COPYRIGHT 2008 Charles Scribner's Sons

(also "**Student** ")

(*b*. Canterbury, England, 13 June 1876; *d*. Beaconsfield, England, 16 October 1937)

statistical theory.

The eldest son of Col. Frederic Gosset and Agnes Sealy, Gosset studied at Winchester College and New College, Oxford. He read mathematics and chemistry and took a first-class degree in natural sciences in 1899. In that year he joined Arthur Guinness and Sons, the brewers, in Dublin. Perceiving the need for more accurate statistical analysis of a variety of processes, from barley production to yeast fermentation, he urged the firm to seek mathematical advice. In 1906 he was therefore sent to work under Karl Pearson at University College, London. In the next few years Gosset made his most notable contributions to statistical theory, publishing under the pseudonym "Student" He remained with Guinness throughout his life, working mostly in Dublin, although he moved to London to take charge of a new brewery in 1935. He married Marjory Surtees in 1906; they had two children.

All of Gosset's theoretical work was prompted by practical problems arising at the brewery. The most famous example is his 1908 paper, "The Probable Error of a Mean," He had to estimate the mean value of some characteristic in a population on the basis of very small samples. The theory for large samples had been worked out from the time of Gauss a century earlier, but when in practice large samples could not be obtained economically, there was no accurate theory of estimation. If an n fold sample gives values $X_1X_2\ldots X_n$ the sample mean

is used to estimate the true mean. How reliable is the estimate? Let it be supposed that the characteristic of interest is normally distributed with unknown mean μ and variance σ^2. The sample variance is

It was usual to take s as an estimate of σ; if it is assumed that $\sigma=s$, then for any error e, the probability that $|m-\mu|\leq e$ can be computed; and thus the reliability of the estimate of the mean can be assessed. But if n is small s is an erratic estimator of σ; and hence the customary measure of accuracy is invalid for small samples.

Gosset analyzed the distribution of the statistic $z=(m - \mu)/s$. This is asymptotically normal as n increases but differs substantially from the normal for small samples. Experimental results m ands map possible values of z onto possible values of μ. Through this mapping a probability, that $|x-\mu|\leq e$ is obtained. In particular for any large probability. Say 95 percent, Gosset could compute an error e such that it is 95 percent probable that $|x - \mu|\leq e$.

R. A. Fisher observed that the derived statistic $t= (n -, 1)½ Z$ can be computed for all n more readily than z can be. What came to be called Student's t-test of statistical hypotheses consists in rejecting a hypothesis if and only if the probability, derived from t. of erroneous rejection is small. In the theory of testing later advanced by Jerzy. Neyman and Egon S. Pearson, Student's t-test is shown to be optimum. In the competing theory of fiducial probability advanced by R. A. Fisher, t is equally central.

Gosset was perhaps lucky that he hit on the statistic which has proved basic for the statistical analysis of the normal distribution. His real insight lies in his observation that the sampling distribution of such statistics is fundamental for inference. In particular, it paved the way for the analysis of variance, which was to occupy such an important place in the next generation of statistical workers.

BIBLIOGRAPHY

Gosset's "*Student's" Collected Papers* were edited by E. S. Pearson and John Wishart (Cambridge-London, 1942; 2nd ed., 1947).

For further biography, consult E. S. Pearson, "Student as Statistician," in *Biometrika*. **30** (1938), 210–250; and "Studies in the History of Probability and Statistics, XVII," *ibid*,**54** (1967), 350–353; and "...XX," *ibid*, **55** (1968). 445–457.

Ian Hacking